a. Subgraph Isomorphism. The Examiner asserted that the specification does not teach or suggest application of subgraph isomorphism for comparison of molecular descriptions as recited in claim 75. In response, Applicants respectfully direct the Examiner's attention to page 30, lines 15-17 of the specification, which teaches that "[a] computer may identify an MCS among a set of molecules by employing subgraph isomorphism, which is a technique well known to those skilled in the art."

b. Genetic Algorithms. The Examiner asserted that the specification does not teach or suggest application of a genetic algorithm for comparison of molecular descriptions as recited in claim 76. In response, Applicants respectfully direct the Examiner's attention to the extensive teaching of use of a genetic algorithm for this purpose, extending from page 31, line 30, to page 33, line 29.

c. Contiguous or Non-Contiguous Molecular Structures. The Examiner asserted that the specification does not teach or suggest finding a molecular feature set that is a contiguous or non-contiguous combination of molecular features. In response, Applicants respectfully direct the Examiner's attention to page 32, lines 1-3 of the specification, which teaches that "[I]n an exemplary embodiment, the maximum common substructure (MCS) should be a contiguous common substructure among the molecules in the hot spot. However, the common substructure may alternatively be a non-contiguous structure."

point out how those two terms are equivalent. In this regard, the Examiner asserted that the term

In response, Applicants respectfully direct the Examiner's attention to page 5, lines 15-23, of the specification as filed, which teaches as follows:

> As used herein, the term "pharmacophore" may mean, without limitation, a representation of any chemical feature or combination of chemical features, including but not limited to features that may be represented in two or three dimensions (e.g., atoms and bonds) and/or other features (e.g., properties associated with the arrangement of atoms and bonds such as proton donors and proton acceptors, electron density space, molecular weight, molecular dipole, etc.). In this sense, the term "pharmacophore" may refer to the mechanism by which molecules in the library interact with a specified target or the mechanism by which molecules evidence any other activity. Further as used herein, the term "structure" may mean, without limitation, a two or three-dimensional arrangement of atom(s) and bond(s) and/or one or more properties conveyed by an arrangement of atom(s) and bond(s).

Further, the Examiner should note that Applicants claims are not directed simply to identifying a molecular feature set in the abstract. Rather, Applicants claims relate to identifying a molecular feature set *that is likely to be responsible for a given observed activity*; in other words, Applicants claims relate to finding a *pharmacophore*, as that term would be clearly understood by those of ordinary skill in the art.

Next, the Examiner noted that Applicants have used the term "identifying" instead of "producing" in claiming the process of discovering a molecular feature set that is likely to be responsible for a given activity, and the Examiner asked Applicants to point out how those two terms are equivalent. Applicants submit that § 112 does not require such equivalence. Rather, at issue is whether the specification teaches the "identifying" limitation, and Applicants submit that it does.

composite structure) accounts for similarities in activity of various molecules" (emphasis

selected group back to the entities within the group, in order to *identify* a subset of common features or components of the entities" (emphasis added)); page 8, lines 22-23 ("adaptively learned substructures may then serve as new descriptors for use in further classifying the molecules in order to *identify* pharmacophoric mechanisms" (emphasis added)).

The Examiner additionally stated that "the extent of what the pharmacophore/molecular feature set is based upon in terms of the instant invention remains unclear." In response, Applicants submit that the specification as filed clearly teaches what is involved in identifying a pharmacophore.

2. Response to Rejections under 35 U.S.C. § 112, Second Paragraph

a. Use of the term "reflecting"

The Examiner next rejected claims 69-90 under 35 U.S.C. § 112, second paragraph, as being indefinite, in that the claims use the word "reflecting" in the context of molecules in a group having "activity characteristics reflecting a given activity." Applicants submit that the meaning of this word is evident from the specification as filed.

In particular, the specification explains that each molecule has a respective activity characteristic that reflects one or more activity levels of the molecule. In accordance with Applicants' invention, after grouping molecules based on structural similarity, a computer seeks to identify "hot spots," in which molecules exhibit a high concentration of a particular activity. (See the specification, at page 26, line 15 - page 29, line 10.) Since the computer looks at the

would be selected based on an extent to which the molecules in the group have activity

Although Applicant believes that the phrase "activity characteristics reflecting the given activity" as used in the claims is clear and understandable to a person of ordinary skill in the art in view of Applicants' description, Applicants have amended the claims to remove the term "activity characteristics reflecting." Thus, the claims now recite, for instance, selecting at least one group of structurally similar molecules based on an extent to which the molecules in the selected group have the given activity." In the exemplary embodiment, this will be done by considering the activity characteristics of the molecules in the group. And by deleting the language "activity characteristics reflecting" from the claims, Applicants do not intend to preclude such consideration of activity characteristics.

b. Use of the term "SOM"

The Examiner also rejected claims 71, 86 and 87 as being vague and indefinite for their use of the acronym "SOM". Applicants submit that the term "SOM" can be interpreted in the way the term is defined in Applicants' specification, and those of ordinary skill in the art would readily understand the term "SOM."

Nevertheless, without changing the meaning of the term in any way, Applicants have amended claims 71 and 86 to recite "Self-Organizing-Map (SOM)" in place of "SOM". The use of the term "SOM" in claim 87 then inherits this definition as set out in claim 86.

c. Use of the term "in response to"

The Examiner next rejected claims 72 and 73 as being vague and indefinite for their use

3. Response to § 102(a) Rejection Over Chen

The Examiner next rejected claims 69, 70, 72-74, 79-84 and 88-90 under 35 U.S.C. § 102(a) as being anticipated by Chen. Applicants respectfully traverse this rejection, because Chen fails to teach every element of the invention as recited by any of these claims as amended. At a minimum, for instance, Chen fails to teach grouping molecules based on similarity of their structural descriptions and without consideration of their respective activity characteristics, and then selecting at least one of the groups based on an extent to which the molecules in the group have a particular activity, in the manner recited in each of Applicants' independent claims.

Chen is a basic example of recursive partitioning (RP), which was described and contrasted in Applicants' priority patent application, U.S. Provisional Patent Application No. 60/120,701, entitled "Artificial Intelligence Directed Lead Discovery," filed February 19, 1999. In the present application, Applicants incorporated the entirety of the priority application by reference.

Applicants' claimed invention is an example of an unsupervised classification method, in which molecules are grouped together based on their structural similarity and *then* at least one of the groups is selected based on an extent to which the molecules in the group exhibit a particular activity. In stark contrast, the RP method described in Chen is a supervised classification, in which molecules are grouped based on their activity into "actives" and "inactives," and then an effort is made to find a descriptor that best distinguishes the actives from the inactives.

[illegible]

[U]nlike the supervised RP system described above, the present invention provides an unsupervised learning technique. The RP system is guided [illegible]

of the data itself. In the preferred embodiment, the present invention does not group compounds by activity (responses to targets) as does the RP system. Rather, the preferred embodiment groups compounds according to their structural similarities.

(Emphasis added.)

Further, Applicants do not find in Chen any teaching of the iterative process set out in claim 84, in which learned substructures are added into a set of molecular substructure keys and then used to classify molecules and so forth.

For at least these reasons, Applicants submit that the invention as recited in claims 69, 70, 72-74, 79-84 and 88-90 patentably distinguishes over the Chen reference.

4. Response to § 102(b) Rejection Over Brown

The Examiner next rejected claims 69, 70, 72-74, 79-84 and 88-90 under 35 U.S.C. § 102(a) as being anticipated by Brown. (The undersigned and the Examiner have determined that the citation to Brown as set out in the office action was in error; it is Applicants' understanding that the Examiner meant to cite Brown at *J. Chem. Inf. Comput. Sci* 1996, 36, 572-584.)

Applicants respectfully traverse the rejection of these claims as being anticipated by Brown, because Brown fails to teach every element of the invention as recited by any of these claims. At a minimum, for instance, Brown fails to teach the combination of functions comprising (i) grouping molecules based on their structural similarity, (ii) selecting at least one of the groups based on the extent to which the molecules in the group exhibit a particular activity, and then (iii) for each group, identifying at least one molecular feature set that is

molecules as recited in Applicants claims. And Brown consequently fails to teach identifying a

In setting forth this rejection, the Examiner asserted that Brown teaches identifying a maximum common substructure. But Applicants disagree. What Brown teaches is determining whether a particular key (2D fragment) appears in molecules and how many times it appears. But the frequency of appearance of a given key does not equate to finding a maximum common substructure.

Consequently, Applicants submit that the invention as recited in claims 69, 70, 72-74, 79-84 and 88-90 patentably distinguishes over the Brown reference.

5. Conclusion

In view of the foregoing, Applicants submit that all of the pending claims are in condition for allowance. And Applicants respectfully request favorable reconsideration and allowance.

6. Information Disclosure Consideration

The Examiner also maintained a refusal to consider some of the references that Applicants submitted in an IDS, because the Examiner asserted the print-out date on the references was not when the information became available to the public.

Applicants respectfully reiterate that the print-out dates on the references is the best information Applicants have. If the Examiner would like Applicants to resubmit the references with the print-out dates redacted, Applicants would be pleased to do so.

Applicants respectfully request the Examiner to consider these references.

Respectfully submitted,

McDONNELL BOEHNEN

Date: November 27, 2002

By: Lawrence H. Aaronson

Marked up Claims Pursuant to 37 C.F.R. § 1.121

69. (Amended) A computerized method of identifying a molecular feature set likely to be responsible for a given activity, based on a set of input data that represents molecules and that defines respectively for each molecule a molecular structure and an activity characteristic, the method comprising:

establishing for each molecule a respective description, by comparison of the molecule's molecular structure to a set of molecular substructure keys;

grouping the molecules based on similarity of their respective descriptions and without consideration of their respective activity characteristics, so as to define groups of structurally similar molecules;

selecting at least one of the groups of structurally similar molecules based on an extent to which the molecules in the selected group have [activity characteristics reflecting] the given activity;

for each of the at least one selected group, identifying at least one molecular feature set common to all of the molecules in the selected group; and

outputting data indicative of at least one identified molecular feature set.

71. (Amended) The computerized method of claim 70, wherein the clustering algorithm comprises [SOM] Self-Organizing-Map (SOM) clustering.

72. (Amended) The computerized method of claim 69, wherein selecting at least one of the groups based on an extent to which the molecules in the selected group have [activity characteristics reflecting] the given activity comprises:

selecting a group [in response to] because the group [containing] contains at least a

73. (Amended) The computerized method of claim 69, wherein selecting at least one of the groups based on an extent to which the molecules in the selected group have [activity characteristics reflecting] the given activity comprises:

selecting a group [in response to] because at least a predetermined percent of the molecules in the group [having an activity characteristic reflecting] have the given activity.

83. (Amended) A method of identifying a molecular feature set likely to be responsible for a given activity, the method comprising:

receiving into a computer a set of input data that represents molecules and that defines, respectively for each molecule, a molecular structure and an activity characteristic;

operating the computer to establish for each molecule a respective description vector, by comparison of the molecule's molecular structure to a set of molecular substructure keys;

operating the computer to apply a clustering algorithm so as to sort the molecules into groups based on similarity of their respective description vectors and without consideration of their respective activity characteristics;

operating the computer to select at least one of the groups based on an extent to which the molecules in the selected group have [activity characteristics reflecting] the given activity;

operating the computer to identify, for each of the at least one selected group, a maximum common substructure of the molecules in the selected group; and

outputting from the computer data indicative of at least one identified molecular feature set.

84. (Amended) A computerized method of identifying a molecular feature set likely to be responsible for a given activity, based on a set of input data that represents molecules and that defines respectively for each molecule a molecular structure and an activity

(b) grouping the molecules based on similarity of their respective descriptions and without consideration of their respective activity characteristics, so as to define groups of structurally similar molecules;

(c) selecting at least one of the groups of structurally similar molecules based on an extent to which the molecules in the selected group have [activity characteristics reflecting] the given activity;

(d) for each of the at least one selected group, identifying at least one molecular feature set common to all of the molecules in selected group;

(e) adding at least one identified molecular feature set as a new substructure key in the set of molecular substructure keys, so as to produce a modified set of molecular substructure keys, and then repeating elements (a) through (d) using the modified set of molecular substructure keys as the set of molecular substructure keys; and

(f) outputting data indicative of at least one identified molecular feature set.

86. (Amended) The method of claim 84, wherein grouping the molecules based on similarity of their respective descriptions comprises [SOM] Self-Organizing-Map (SOM) clustering the molecules based on their respective descriptions.

90. (Amended) A processing system for modeling chemical structure-activity relationships through artificial intelligence analysis of an input data set representing molecules, each of the molecules having a set of features and an activity characteristic, the processing system comprising, in combination:

means for establishing for each molecule a respective description, by comparison of the molecule's molecular structure to a set of molecular substructure keys;

means for grouping the molecules based on similarity of their respective descriptions and without consideration of their respective activity characteristics, so as to define groups of

an extent to which the molecules in the selected group have [activity characteristics reflecting] the given activity;

means for identifying at least one molecular feature set common to all of the molecules in each of at least one selected group; and

means for outputting data indicative of at least one identified molecular feature set.